DOCUMENT RESUME

ED 413 328 SP 037 639

AUTHOR Mertler, Craig A.
TITLE Students as Stakeholders in Teacher Evaluation: Teacher Perceptions of a Formative Feedback Model.
PUB DATE 1997-10-00
NOTE 31p.; Paper presented at the Annual Meeting of the Mid-Western Educational Research Association (Chicago, IL, October 15-18, 1997).
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Evaluation Methods; Feedback; Formative Evaluation; High Schools; Higher Education; Models; *Student Attitudes; *Student Evaluation of Teacher Performance; *Teacher Attitudes; *Teacher Student Relationship
IDENTIFIERS Florida

ABSTRACT

The "Student Evaluation of Teaching and Teaching Techniques" (SE3T) instrument, designed for this study of student evaluation of teacher performance, was administered twice to students in seven high schools in north Florida. Each time, results were tabulated and feedback provided to each teacher within one week; each teacher received only his or her own ratings and comments. A survey on the perceived strengths and weaknesses of the process was distributed to participating teachers. Analyses of the data revealed that overall reaction of the teachers to this process of collecting student feedback was highly positive and that it was feasible for teachers to conduct this process on their own. Teachers expressed concern, however, that the process would only be useful if students were honest in their ratings and comments, and that enough time were to elapse between administrations of the SE3T for teachers to process and act upon the feedback. Nearly all of the teachers discussed behaviors, identified as weaknesses by their students, that they had already addressed or were planning to address in the future. Results of the study suggested that the process of collecting student feedback on teachers' classroom performances is useful to teachers who voluntarily participate in the process, and that the SE3T rating form is a useful and valid instrument for collecting the data. Two data tables and a diagram of the SE3T model are appended. (Contains 39 references.) (ND)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 413 328

Running Head: TEACHER PERCEPTIONS OF FORMATIVE FEEDBACK

STUDENTS AS STAKEHOLDERS IN TEACHER EVALUATION: TEACHER PERCEPTIONS OF A FORMATIVE FEEDBACK MODEL

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

C. Mertler

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Craig A. Mertler
Bowling Green State University
Bowling Green, OH 43403

Paper presented at the
Annual Meeting of the Mid-Western Educational Research Association
Chicago, IL
October 15-18, 1997

SP 037639

BEST COPY AVAILABLE

INTRODUCTION

The Importance of Sound Evaluation of Classroom Teaching

Virtually all educational institutions in the United States evaluate the qualifications and work of their personnel. These evaluations occur at several key "points" during an individual's period of service with the institution, including certification, selection (hiring), assignment, promotion, award of tenure, and allocation of special recognition or awards (Joint Committee on Standards for Educational Evaluation, 1988). Many institutions also use evaluation as a means to provide feedback for improving the performance of educational personnel. The need for sound, thorough evaluation of educational personnel should be clear. As stated in the Personnel Evaluation Standards (Joint Committee on Standards for Educational Evaluation, 1988), " . . . to educate students effectively and to achieve other related goals, educational institutions must use evaluation to select, retain, and develop qualified personnel and to manage and facilitate their work" (p. 5). It is evident that personnel evaluation in education has been used to select and retain teachers, but seldom has it been used for the development of qualified teachers.

Historically, there has been widespread dissatisfaction with the quality of personnel evaluation in education (Joint Committee on Standards for Educational Evaluation, 1988). Educators, policy makers, and community groups often attack the near absence of personnel evaluation systems or the superficiality in the systems that do exist (p. 157). Highly developed and effective teacher evaluation systems are rare in American education (Darling-Hammond, 1986). Darling-Hammond, Wise, and Pease (1983) state that "most existing systems [for evaluating teachers] are illogical, simplistic, unfair, counterproductive, or simply unproductive" (p. 158). In their review of procedures used to evaluate teachers, Soar, Medley, and Coker (1983) agree that these evaluation procedures possess all of these characteristics. Frase and Streshly (1994) assert that "teacher . . . evaluation appears to be purely ceremonial, with little or no intent to improve instruction . . ." (p. 50). Finally, Scriven (1980) declares that the procedures used in the evaluation of teaching are "shoddy at the intellectual and the practical levels" (p. 1).

Teacher evaluation is just one of many aspects in the lives of teachers. It is an integral component of the profession, from preservice training through certification, contract renewal, and professional development (Nevo, 1994). However, it has come to be viewed by those it was meant to help as a means of controlling teachers, motivating them, holding them accountable for their performance, or getting rid of them when their performance is poor. Teachers' resistance to evaluation is reasonable if the evaluation "is subjective, unreliable, open to bias, closed to public scrutiny, and based on irrelevancies" (Soar, Medley, and Coker, 1983, p. 246).

BEST COPY AVAILABLE

The vast majority of personnel evaluation in education occurs at the summative level. Summative evaluations are used almost exclusively for personnel decisions (i.e., for hiring, determining promotions, dismissal of a teacher, etc.) (Haefele, 1993). These evaluations or performance reviews are characteristically conducted by an administrator typically untrained in personnel evaluation (Levin, 1979). Furthermore, these evaluations suffer from a serious lack of adequate sample size — often 40 to 60 minutes of observation per school year (Haefele, 1993; Darling-Hammond, 1986; Savage & McCord, 1986).

Serious problems exist when this method is relied upon as the sole source of information regarding the evaluation of teacher performance and, unfortunately, this distressing situation is all too common in American schools. This reliance on a single source of information may result in the certification of unqualified persons, a lack of constructive feedback to individual teachers, or failure to recognize outstanding service, among other consequences. For these reasons, many concerned and informed groups agree that there is a serious need both to increase the amount and improve the quality of personnel evaluation in education. Many state education departments, school districts, and other policymaking groups have moved to devise better systems of personnel evaluation, which have included programs such as merit pay, career ladders, professional certification tests, peer evaluation, and master teacher status (Joint Committee on Standards for Educational Evaluation, 1988).

Why Formative Strategies for Evaluating Classroom Teaching?

An alternative which is quickly gaining in popularity is the inclusion of formative personnel evaluation into systems of teacher evaluation. Formative personnel evaluation is designed to help teachers improve the quality of instruction by identifying strengths and weaknesses (Haefele, 1993). The improvement of instructional practice is quite possibly the most important — and most positive — purpose of teacher evaluation (Manning, 1988). However, this method of teacher evaluation continues to be extremely under-utilized (Haefele, 1980).

In recent years, increasing numbers of teachers have begun to show interest on their own in using evaluation techniques for self-evaluation purposes in order to improve their own teaching performance (Nevo, 1994). It is the intent of this form of evaluation to help teachers improve their teaching performance according to the needs of their *own* students, and hence improve students' learning. These techniques include peer observation, self-assessment activities, and student ratings of teachers' classroom performance, among others. Teachers should be encouraged to take the initiative in using evaluation for self-improvement. The alternative for teachers is simply the *status quo* — waiting for their principals or school districts to impose evaluation on them.

THE NEED FOR MULTIPLE PERSPECTIVES IN TEACHER EVALUATION

Failure of the "Dominant Model" of Teacher Evaluation

A standard model of evaluating teaching has been employed in schools for decades (Haefele, 1993). This model continues to be the most widely used and accepted model in the public school systems (Bailey, 1978; Levin, 1979). It is truly the "dominant model" of teacher evaluation.

In most school districts, nontenured teachers are observed two to three times each year (Gitlin & Smyth, 1990; Haefele, 1993; Levin, 1979; Savage & McCord, 1986). The evaluations are usually conducted by the building principal (who, according to the stereotype, stands "stonefaced at the back of the classroom filling out a form") and last between 20 and 30 minutes (Darling-Hammond, 1986, p. 531; Haefele, 1993; Johnston & Hodge, 1981). Tenured teachers are observed for approximately the same length of time, but less frequently (Haefele, 1993). Sometimes, the observation is followed by a conference, where the principal's responsibility is to describe the strengths and weaknesses of the teacher's performance and offer suggestions for improvement (Haefele, 1993). This process typically involves one-way communication from the evaluator to the teacher (Gitlin & Smyth, 1990). All teachers are evaluated using a common instrument and fixed procedures.

Advantages of this dominant model include: (a) Administrators who possess skill and expertise can significantly assist teachers who desire to improve their instructional performance, (b) the administrator can demonstrate his or her concern for quality instruction to his entire staff, and (c) evaluation, retention, tenure, and salary increments may be determined with this model (Bailey, 1978).

However, research has shown that even this deficient model is not employed in a thorough, consistent manner. Huddle (1985) summarizes selected results from the National Institute of Education's "High School and Beyond Teacher and Administrator Survey." Teachers were asked how often they were observed by *any* supervisor during the previous year. One-fourth (26%) of the respondents indicated "never," another one-fourth (27%) responded "only once," and nearly another one-fourth (23%) indicated "twice." Infrequently performed evaluation, such as those experienced by the teachers responding to this survey, certainly have limited validity (Huddle, 1985).

Darling-Hammond (1986) describes the dominant model of teacher evaluation as "a perfunctory bureaucratic requirement that yields little help for teachers and little information on which a school district can base decisions" (p. 531). This standardized process serves "ideally" as the all-inclusive means for discussions of individual teaching practice, for guidance regarding professional development, and for personnel decision making (Darling-Hammond, 1986). Scriven (1990) has described the dominant model of teacher evaluation as nothing more than reports from occasional visitors to the classroom. These reports, Scriven

says, "suffer from samples that are inadequate in size and not representative, measurement artifacts, style bias, and failures of empathy, and are usually vulnerable to personal bias" (p. 91).

Veir (1990) recounts an even more dismal view of the dominant model. Her scenario begins with the teacher being summoned to the principal's office where the principal proceeds to complete a blank evaluation form with little or no evidence and little direct experience. There is no feedback for the teacher and no suggestions for improvements (p. 7). These evaluative exercises do little more for teachers than add to their cynicism of a bureaucratic routine.

These views emphasize some of the ways in which this model is less effective, specifically in the area of instructional improvement (Bailey, 1978). The disadvantages of the dominant model are numerous and include: (a) The teacher is almost totally dependent on the evaluator to collect and analyze information, (b) the quality of improvement of instruction is closely aligned with the accuracy of a single evaluator's perceptions, (c) the evaluator seldom takes the time to share methods, processes, or techniques that could assist the teacher in correcting weaknesses and maintaining that behavior, and (d) teachers often perceive these evaluation activities as a threat (Bailey, 1978, p. 54).

Other weaknesses of this model address the fact that most evaluators (principals) have had no training in observational techniques and do little or no preparation prior to observing a teacher (Levin, 1979). Cangelosi (1991) claims that unless observers are trained to use well-designed observational instruments, observations of teachers in classrooms "will continue to be dominated by malpractice that produces invalid results" (p. 47). Principals have little time for the evaluation of teaching personnel. They may be unwilling, or simply unable, to devote the time necessary to conduct thorough evaluations (Savage, 1982). It is unreasonable to expect them to have specialized content area or pedagogical knowledge of *all* teaching areas in which they are required to evaluate staff. Because of these difficulties, many principals do harm to the evaluation task by omitting preobservation conferences, making brief unannounced visits to the classroom, and minimizing the follow-up activity by simply placing a copy of the completed rating form in the teacher's mailbox, as if it were a "report card" (Savage, 1982, pp. 41-42). Needless to say, this report card style of teacher evaluation results in little improvement in teaching effectiveness (Jacobson, 1986).

Finally, this dominant form of teacher evaluation serves to separate teachers from one another. It does not encourage them to establish a communicative environment where knowledge and practices are shared. Rather, it invites them to compete against each other so as to look better on the evaluative criteria (Gitlin & Smyth, 1990).

The standard model of teacher evaluation used in American schools is in serious trouble (Haefele, 1993). There are several reasons for making this claim. Evaluation criteria lack validity; evaluators are untrained; weak and incompetent teachers are consistently

awarded lenient ratings (p. 21). Many school districts lack clearly stated purposes for the evaluation of teaching. These ineffective practices have permitted unqualified persons to assume teaching positions, have made it difficult to rid education systems of incompetent and unproductive teachers, and have failed to provide direction for staff development (Haefele, 1992; Haefele, 1993; Joint Committee on Standards for Educational Evaluation, 1988). Teachers often view these forms of evaluation as frivolous, unfair, and inaccurate (Johnston & Hodge, 1981).

As an evaluation system, the dominant model is lacking three essential components, previously eluded to in this section. These include: (a) Evaluators lack two essential sets of skills needed to evaluate teachers: skills in evaluating teacher performance, and skills in communicating with teachers about the results of the evaluation process; (b) there is insufficient time for both evaluations and follow-up activities; and (c) the process for linking staff development and teacher evaluation is not clear or does not exist (Stiggins & Duke, 1988).

The dominant model of teacher evaluation simply does not accomplish the goal of providing teachers with a means to effectively improve their instructional practices. Unfortunately, the individuals ultimately affected by insufficient methods of instructional evaluation are the students. If teachers are not provided with feedback to improve their instruction, students cannot benefit from that improved instruction. As educators, if we want to improve teacher performance and the relationship between teacher performance and student outcomes, continued reliance on the dominant model of teacher evaluation won't accomplish it; this model is simply ineffective as a process for improving the performance of teachers (Haefele, 1992). We must promote methods of evaluation which cultivate the improvement of instruction and we must do so by using alternative sources of information.

Teachers as Stakeholders in Teacher Evaluation

Scriven (1991) defines formative evaluation, in general, as evaluations that are "typically conducted during the development or improvement of a program or product (or person, and so on) and it is conducted, often more than once, *for* the in-house staff of the program *with the intent to improve*" (pp. 168-169). He specifically defines formative teacher evaluation as that which is done to assist professional development (Scriven, 1994).

Teachers have a more favorable attitude toward evaluation when the results are used in a formative manner to help them improve performance (Haefele, 1993). The individual teacher typically decides on the evaluation criteria on the basis of personal interest, by requesting feedback in the areas of their teaching they would like to improve (McColskey & Egelson, 1993).

BEST COPY AVAILABLE

An evaluation is formative if its sole purpose is to provide information that is useful for making decisions about how to teach (Cangelosi, 1991). Teachers engage in formative evaluation whenever they evaluate their own instruction for guidance in organizing, designing, or planning lessons (Cangelosi, 1991). Formative teacher evaluation is, by nature, typically more descriptive and less judgmental (Nevo, 1994). These evaluations can be based on self-evaluation, peer evaluation, or evaluations by principals, students, or parents.

An important attribute of formative teacher evaluation is a more active level of participation by the teacher being evaluated than is evident in summative evaluation systems (Bacharach, Conley, & Shedd, 1990). Teacher involvement in the evaluation process creates a sense of ownership and improves the genuineness, focus, and meaning of the evaluation activity (Huddle, 1985). Barber (1990) characterizes formative teacher evaluation as

> . . . a set of procedures designed to assist teachers in improving their own teaching. Formative teacher evaluation can help an ineffective teacher become a better teacher or an effective teacher become an excellent teacher. Formative evaluation is a helping, caring process that provides data to teachers *for making decisions about how they can best improve their own teaching techniques, styles, or strategies* [italics added]. Formative evaluation must occur in close collaboration with the person being evaluated — he or she must *agree* to it, be an *intensive part* of it, *participate willingly* in it, and, in the case of experienced teachers, even *direct* it [italics added] . . . (p. 216).

Stiggins (1986) further emphasizes the idea of teacher involvement in a formative system. He states that the feedback may come from a variety of sources (e.g., supervisors, peers, students, self-analysis) and may be continuous or intermittent, but it is the teachers' job to (a) evaluate the feedback and (b) take responsibility for acting on that feedback.

The key to a formative evaluation system is that the environment for collecting information is not judgmental or punitive, but rather supportive of growth and teacher-directed (McColskey & Egelson, 1993). It must be nonthreatening to the receiving teacher (Manning, 1988). A formative system can help to (a) encourage continual teacher self-evaluation and reflection and discourage the development of teaching routines that never change, (b) encourage individual professional growth in areas of interest to the teacher, (c) improve teacher morale and motivation by treating the teacher as a professional who is in charge of his or her own professional growth, (d) encourage teacher collegiality and discussion about practices among peers in a school, and (e) support teachers as they try new instructional approaches (McColskey & Egelson, 1993, p. 2).

Students as Stakeholders in Teacher Evaluation

Student evaluation of college instructors has been occurring since the 1920s, when Harvard students actually published their assessments of their teachers' effectiveness (Savage, 1982), and the first published form for collecting student ratings, the *Purdue Rating Scale of*

Instruction, was released in 1926 (Cook, 1989). Their use became widespread during the 1960s (Levin, 1979).

A substantially lesser amount of research has involved the evaluations of secondary (i.e., grades 9 through 12) teachers by their students (Smith & Brown, 1976; Traugh & Duell, 1980) and even less "real" application of this method occurs in high schools (Levin, 1979). Hanna, Hoyt, and Aubrecht (1983) state that student evaluations of teaching effectiveness at the high school level have been "largely neglected."

Students themselves have taken an active role in attempting to motivate and assist faculty in the improvement of classroom instruction at the high school level. The New York State Education Department Student Advisory Committee, comprised of 17 secondary students representing geographic regions across the state, have urged students to make it their responsibility "to share their thoughts, comments, and criticisms not only among their peers, but with the teachers as well" (New York Education Department, 1986, p. 10). They have also recommended the use of student evaluations for the purposes of determining those qualified to receive special recognition. The areas of focus should include, among others, the quality of a prepared lesson, fairness in grading, promptness to class, and general attitude of the teacher. Forms of commendation might include a Teacher Recognition Day or a Teacher of the Month Award.

The Pennsylvania Student Advisory Board also encouraged the use of the *Student Observation of Teachers and Teaching Techniques* (*StOTT*) instrument by teachers in their state during the late 1970s (Pennsylvania State Department of Education, 1981). Their recommendations included the fact that the *StOTT* should not be used to evaluate or compare high school teachers; it is a valuable instrument to be used by teachers in their own classrooms. Furthermore, they suggested that teachers use the form on a purely voluntary basis.

Programs of this type are rare. Barsalou, Killinger, and Thompson (1974) cite a survey of evaluation practices of 213 school districts. In the resultant data, no use of student data is mentioned.

However, student evaluation data can be used for assigning value to the quality of teaching or it can be used to determine specific strengths and weaknesses which can then be used by the teacher to remove and/or retain individual instructional behaviors (Bailey, 1978). They can also be used to evaluate curriculum materials and the teachers' implementation of these materials (Traugh & Duell, 1980). The use of student ratings in teacher evaluation assumes that: (a) the student knows when he or she has been motivated, (b) it is the student whose behavior is to be changed, (c) student rating is feedback to the teacher, and (d) student recognition may actually motivate good teaching (Darling-Hammond et al., 1983).

It has been assumed that since students are the only individuals who are constantly exposed to the various elements of a course (e.g., instructor, textbook, homework, course

content, method of instruction, etc.), they are the most logical evaluators of the effectiveness of those course elements (Aleomoni & Spencer, 1973). They can provide insights that no one else can (Stiggins, 1986). Students are the only observers who are in class on a regular basis; they are in the classroom every day (Jacobson, 1986; Savage & McCord, 1986; Stiggins & Duke, 1988). For example, no one is in a better position to critique the clarity of teacher directions than the students for whom the directions are intended (Stiggins & Duke, 1988). Students are best able to provide feedback about when they are bored or when they are lost or confused (McColskey & Egelson, 1993). Moreover, feedback provided by students fosters the emotional relationship between teacher and student (Tacke & Hofer, 1979). Student ratings have also been shown to provide useful information in identifying teaching behaviors of student teachers perceived by students to be most effective and least effective (Martin, 1988).

Furthermore, it can be argued that a teacher can hardly keep from being influenced by the informal feedback provided by students (Cangelosi, 1991). Students believe that their opinions about teachers do make a difference in the ways teachers teach and that teacher ratings are not a waste of time (Traugh & Duell, 1980). Therefore, it seems reasonable to assume that students take their ratings seriously and mark the rating forms as accurately as they can. Provided the students are not asked to comment on aspects of teaching in which they have no expertise, feedback from them constitutes an excellent data source for professional development (Stiggins & Duke, 1988). It is best to have students evaluate only interactive aspects of teaching, not aspects such as planning (Savage & McCord, 1986).

Advantages of using student evaluations are numerous. They can be an effective means of helping teachers improve their instructional skills (Bailey, 1978). Feedback from students can be collected quickly and at a point in time convenient to all participants (Savage, 1982). Student evaluations are available to teachers whenever they want to use them and they can be an ongoing process that is not dependent upon the assistance of a principal or supervisor (Halbert, 1975). Students become more positive about their learning since they sense that the teacher wants to improve his or her instruction (Bailey, 1978). Additionally, data of this type cannot be labeled as "unrepresentative" (Savage, 1982; Savage & McCord, 1986).

Yet, there are disadvantages to this method of gathering data on teacher effectiveness. This method is based on the assumption that students are being candid and honest in expressing their opinions (Bailey, 1978). If students fear teacher retaliation, the activity becomes relatively ineffective. The teacher must be willing to use student feedback and must demonstrate this willingness (Bailey, 1978). If teachers are not, students will be discouraged from being honest. Total reliance on student feedback to improve instruction could potentially result in the creation of a teacher that *students* believe to be most effective (Bailey, 1978; Owens, 1976).

OVERVIEW OF THE *STUDENT EVALUATION OF TEACHING* STUDY

Design Phase

The selection of subjects for this study came from two school districts in the central portion of north Florida. The individual subjects representing seven high schools constituted a voluntary sample.

The initial draft of the rating instrument to be used in the study was developed by the researcher, following a review of research publications and articles which contained rating instruments of teacher effectiveness, teacher competence, and teacher behaviors. Items which were deemed appropriate for the instrument were selected from several sources in the literature. This technique of item selection has been employed previously in the development of student evaluation forms (Savage & McCord, 1986).

The *Student Evaluation of Teaching and Teaching Techniques* (*SE3T*) instrument was comprised of a total of 37 items. The first 33 items appearing on the *SE3T* were specific teacher behaviors written as individual statements. The students were asked indicate the extent of their agreement with each statement on a four-point Likert-type scale. The definitions of the points included on the scale were: "strongly disagree," "disagree," "agree," and "strongly agree."

Items 34 and 35 were summative-type items. Item 34 asked the student to agree or disagree with a statement regarding a recommendation of this teacher to other students. Item 35 asked the student to rate the overall teaching ability of the particular teacher. Students also responded to this item on a four-point scale, with points defined as follows: "excellent," "good," "average," and "below average."

Items 36 and 37 were open-ended items, giving each student an opportunity to provide comments of a more personal nature. Item 36 asked the student to list one or two things that he or she especially liked about the teacher. Item 37 asked the student to list one or two things that he or she disliked about the teacher.

It is crucial to note that once the teachers volunteered to participate in the study, they were given opportunities to review the draft of the *SE3T* instrument for purposes of providing suggestions or revisions to any items that they did not like or agree with, or with which they anticipated problems. It was important to the researcher that the teachers did *not* feel that they were forced to use an instrument without having an opportunity to provide input into its development.

The rating instrument was designed to be administered to students in grades 9 through 12. In an effort to ensure that the *SE3T* was developed at an appropriate reading level for the students, a readability test was conducted on the instrument using a grammar-checking program contained in Microsoft Word (version 5.1) for the Macintosh computer. The readability procedure resulted in a Flesch Grade Level equivalent to 7.8 (i.e., the eighth

month of grade seven). It was concluded by the researcher, based on comments from the participating teachers, that the language and vocabulary used in the *SE3T* would be appropriate for those students for whom it was intended.

Subsequent to review of the *SE3T* instrument by the teachers involved in the study, and making appropriate revisions, a pilot test of the instrument was conducted. The purpose of this pilot study was to enable the researcher to determine how well the items were understood and to check for the presence of ambiguous terms.

The pilot test was conducted with approximately 40 high school students who did *not* participate in the regular study. Following the administration of the *SE3T* instrument, the teachers asked the students if they experienced any difficulties in reading or understanding any of the items. The students were also asked to offer suggestions for improvements which could be made to the instrument.

Data Collection Phase

The *SE3T* instrument was administered to the students on two occasions during the fall of 1995-1996 school year. The teachers conducted the first administration during approximately the seventh week of the school year (i.e., early October). The second administration was conducted approximately six weeks later (i.e., early to mid-November).

The rating forms were collected by a designated student within each class for each teacher. The method used for the collection of the data was as follows. A designated or volunteer student within each class distributed one rating form and one plain white envelope to each student. Students were then asked to complete the form without placing their names anywhere on the form, fold it, place it in the white envelope, and seal the envelope. The designated student then collected the individual white envelopes and placed them in a large manila envelope, which was sealed in the presence of the students and placed in a box located in the main office. (In some cases, the teacher simply placed the manila envelope on his or her desk, and then instructed the students to seal their white envelopes and come to the front of the room to place their individual envelope in the manila envelope.) The researcher then collected the manila envelopes containing the forms from each school.

Following each of the administrations of the *SE3T* instrument to the students, results were tabulated and feedback provided to each teacher within one week. This feedback report consisted of a form very similar to the *SE3T*, in that it contained each item on the instrument grouped according to the appropriate domain of teacher behavior. Specifically, the feedback included the mean rating for each item and the combined positive (i.e., "agree" plus "strongly agree") and negative (i.e., "disagree" plus "strongly disagree") responses, reported as percentages. Finally, each teacher was provided with a list of open-ended comments supplied by his or her students. It is important to recognize that there existed the potential for these comments to affect student anonymity. This potential was minimized as a

result of the researcher supplying only *typed* copies of the comments to the teachers; they did not see the actual handwritten comments. In addition, the researcher chose to delete certain student comments that were of a derogatory nature or that contained profanity. Teachers were informed upon receipt of their individual feedback reports that selected comments from their students may have been deleted for those reasons.

It is important to note that each teacher received only his or her specific ratings and comments. Data and other information resulting from the analysis of student evaluations of individual teachers remained the property of the individual teachers, for their own review and processing. Teachers were not be provided with summary information for any other teacher involved in the study.

Assessment Phase

Two sources of information were used to assess the evaluation process: (1) analyses of the ratings and responses provided by the teachers to questions on a follow-up survey and (2) analysis of the notes and transcription of a focus group session conducted with a subgroup of the participating teachers.

ASSESSMENT OF THE STUDENT FEEDBACK PROCESS

Teacher Feedback Concerning the Process

Approximately one week following the second administration of the *SE3T* form, participating teachers were surveyed in order to gather some initial feedback on this process. The purpose of the *Teacher Survey* was to offer *all* teachers the opportunity to provide feedback and discuss their perceptions of the strengths and weaknesses of this process of receiving student feedback. The *Teacher Survey* was given to teachers at the time they received the results of the second administration of the student rating form. Pre-addressed and stamped return envelopes were also provided. The teachers were given two weeks to return the surveys.

The survey consisted of eight items. Items 1 through 4 were forced-choice items, asking the teachers to rate the usefulness of the process of receiving student feedback, the appropriateness of the items appearing on the *SE3T* form, the feasibility of collecting student feedback in this manner on their own, and how often student feedback should be collected. Additional space for comments was provided for each of these items.

The remaining four items were open-ended. Item 5 asked the teachers what, if any, changes they had made to their teaching. Item 6 asked the teachers to explain what was most useful about this process of collecting student feedback and Item 7 asked teachers what was

least useful about this process. Finally, Item 8 asked the teachers to explain how this process could be improved in order to be more useful.

Approximately three weeks following the administration of the *Teacher Survey* (i.e., early December), 5 to 8 participating teachers were sought for participation in a focus group interview scheduled to last approximately 90 minutes. The researcher attempted to secure the participation of a minimum of one teacher from each of the participating high schools (i.e., five teachers).

The purpose of the focus group was to gather in greater depth the personal insights, experiences, and perceptions of the teachers who were evaluated by their students. Additionally, this was a means of conducting a triangulation of the data resulting from the *Teacher Survey*, for the purposes of checking the trustworthiness of the data. It was the intent of the researcher that the information generated from the focus group interview would reveal more about the strengths and weaknesses of the process of providing teachers with feedback from their students than was gathered as a result of the *Teacher Survey*.

Aspects of the Process Examined During the Assessment Phase

Most of the teachers shared their excitement about participating from the outset of the study. The teachers shared comments with the researcher such as "I can't wait to see the results!" Some teachers discussed their initial concerns and hesitancy, which were eventually quelled by their participation and the participation of their students.

Teachers were asked to rate aspects of the student feedback instrumentation (*SE3T*) and the procedures used to collect the feedback. The analysis of the responses provided by the teachers was conducted using SPSS-PC+ (version 4.0.1). A summary of the frequencies and percentages of responses to the four scaled items is presented in Table 1 (see Appendix A). Additionally, comments supplied by the teachers on the *Teacher Survey* and in the focus group session, and identified as strengths and concerns of the process, are summarized in Table 2 (see Appendix B).

Usefulness of the Process

The overall reaction of the teachers to this process of collecting student feedback on their teaching was highly positive. As shown in Table 1, all of the teachers responding to the *Teacher Survey* rated the overall process as either "extremely useful" or "somewhat useful".

Strengths. The positive ratings of the usefulness of the process were supported by many additional comments supplied by the teachers regarding why or how it was useful. Some stated that it gave them an opportunity to see themselves from another point of view. Teachers also believed that the process was important to and enjoyed by the students, giving them a chance to voice their opinions. One teacher seemed pleasantly surprised by the

sometimes insightful comments she received from her students. She expressed shock and disbelief after reading many of these unexpected comments. This teacher, in turn, seemed to appreciate her students even more.

The process seemed to be especially useful for one teacher, who had taught high school juniors and seniors for the past 27 years and was now teaching at the seventh grade level. Having grown accustomed to teaching students four and five years older than her current students, she believed that receiving feedback from the seventh grade students was crucial to her ability to adapt her methods to the younger students and, thus, experience success in working with them.

Interestingly, less experienced teachers also found the experience to be an invaluable one. The process was important in revealing areas in need of improvement early in one's professional teaching career.

Finally, several teachers commented in a highly positive manner on the method of collecting the student rating form specifically in an effort to maintain the students' anonymity. They believed this procedure ensured a higher level of student honesty in their responses.

Concerns. Although the overriding sentiment regarding the usefulness of the process was positive, teachers also shared their concerns about ways in which the process was less useful. They offered many of these in the form of cautionary statements. They believed that the process would only be useful if the students were being honest in their ratings and comments. It was also implied that the process is useful only to the extent a teacher changes his or her behavior based on the student feedback.

The amount of time between the first and second administrations was also seen as a contributing factor to the non-utility of the process. Teachers stated that there was not ample time to process and act upon the feedback, nor was all of the necessary and required information to do so provided.

Appropriateness of *SE3T* Items

As shown in Table 1, when teachers were asked to rate the appropriateness of the items appearing on the *SE3T* form, 10 (83%) out of 12 responded "extremely appropriate" and only two (17%) responded "somewhat appropriate".

Strengths. Teachers commented that the *SE3T* rating form covered a broad range of teacher behaviors, yet remained quite relevant. Teachers believed that both sections of the student rating form — the scaled items and the open-ended items — were equally useful. They did not see one as being more beneficial than the other; both served a specific purpose. Furthermore, they thought that neither section could sufficiently stand alone as a unique student feedback instrument.

Additionally, the teachers were asked if they intended to continue using the *SE3T* or similar type of student rating form. Many indicated that they would use this form because they liked the items which comprised it. Those who had not solicited evaluations from their students prior to this experience became aware of beneficial aspects of collecting student opinions.

Concerns. Only a few teachers commented on what they believed to be "problematic items" on the rating form. One teacher had difficulty herself with Item 33 ("My teacher varies the way he/she teaches"). She thought *her* lack of understanding of the item would only be compounded in her students. Although the teacher had some difficulty with this item, she claimed that her students appeared not to be confused.

Feasibility of the Process

Three-fourths (75%) of the teachers responding to the *Teacher Survey* indicated that they believed this process of gathering student feedback was feasible for teachers to conduct on their own (see Table 1). Only 2 teachers responded that the process was not feasible and 1 indicated that he was unsure of its feasibility.

Strengths. Several teachers stated that the time element involved in collecting this information was very reasonable.

Concerns. However, other teachers indicated that a lack of feasibility might be caused mainly by two aspects of the process — the actual administration of the rating form and the compilation of the results. First, teachers suggested the possibility of a having a neutral party monitor the administration of the student rating form and the subsequent collection of those forms. The reason for doing so would be to further protect the anonymity of the students, in addition to fostering honesty in the student responses.

Secondly, teachers expressed concern over the compilation of the results, especially tallying the responses to the scaled items. Two teachers suggested the use of computerized response sheets as a means of partially alleviating this problem.

Frequency of *SE3T* Administration

The majority (64%) of teachers indicated that "twice per school year" was an appropriate number of times to administer the rating form and receive feedback from students; 2 teachers responded that "once per school year" would be appropriate and 2 also indicated that "four times per school year" would be suitable. The teachers declared that they would survey their students once or twice per course. In other words, whether a course lasted a semester or a full school year, the teacher would likely administer the rating form twice during the course.

The two teachers who responded "four times per school year" taught at a school which currently operated on a block schedule; in other words, courses lasted approximately

nine weeks. Each nine weeks, teachers would have a completely new set of students in each course they taught. These teachers said that they would survey their students once per course.

Nearly all of the teachers believed that the two administrations of the *SE3T* rating form were scheduled too close together. One teacher stated that "the amount of time that was given to change or improve teaching methods between evaluations" made the process less useful.

Nature of the Teacher Feedback

Strengths. Teachers were asked to comment on the nature of the feedback they received from the researcher, consisting of the aggregation of the student ratings and comments. Most of the positive comments centered on the students' responses to the open-ended items which appeared on the rating form and the way in which they were reported to the teachers.

Concerns. In contrast, two teachers offered cautionary comments regarding these student-developed responses and how they were presented in the teacher feedback reports.

Several suggestions were offered as ways to improve the teacher feedback reports. These consisted of some minor revisions to the format of the feedback report (i.e., reporting the percentages of responses for all points on the rating scale instead of collapsing categories), as well as some additional information supplied to the teachers. This additional information might consist of differences in the ratings from the first administration to the second (e.g., difference scores) or a list of ways to work on areas in need of improvement.

Changes in Teacher Behavior

Teachers were also asked to describe ways in which they had changed, or were considering making changes to, their teaching behaviors as a direct result of the feedback from their students. Nearly all of the teachers described how they were addressing weaknesses identified by their students or changes they had already made to certain behaviors. These behaviors included teacher empathy toward students, verbal and nonverbal communication skills, pre-instructional planning, and variation in methods of instruction.

Only a couple of comments indicated little or no attempt at improving teaching techniques on the part of teachers.

Several teachers also described behaviors and techniques that they were planning on changing in the future. Most of these addressed ways to make instruction more interesting to students, which these teachers said is something that "is just going to take time."

Teacher Follow-Up to the Process

As a further indication of the earnestness with which teachers bought into this process and the degree to which they processed and followed up on the feedback, several

participating teachers described ways in which they pursued, beyond the basic feedback, improvements to their classroom teaching. They discussed the results with their students in an open forum, soliciting suggestions for improvements. They saw this self-initiated follow-up activity as being very beneficial to them.

Summary

Comments such as those provided above served as evidence that this group of volunteer teachers valued the experience of participating in this study and having the opportunity to collect student feedback on their teaching. For the majority, it was their first experience with gathering student opinion; for others, it served as an opportunity to use a different type of student feedback instrument. In either case, most teachers realized a variety of benefits resulting from the collection of feedback from their students, and many expressed a desire to continue giving students a voice in the evaluation of their teaching. Nearly all of the teachers described behaviors, identified as weaknesses by their students, that they had already addressed or were planning to address in the future. Several concerns were shared and suggestions for ways to improve the process were also offered by the teachers.

CONCLUSIONS AND RECOMMENDATIONS

Many of the findings of this study strongly support both the findings of previous research studies and recommendations of experts in the area of formative teacher evaluation. This process has been shown to be beneficial both to teachers and their students. Additionally, the instrumentation developed for use in this study, with a large degree of teacher input, was shown to be highly useful for its intended purpose.

Specifically, it can be concluded that the process of collecting student feedback regarding teachers' classroom performance is a very useful one to teachers who voluntarily participated in this process. In contrast to the deficit model of teacher evaluation (i.e., infrequent classroom visits and evaluations by administrators), this process is seen as something done *for* teachers, rather than *against* them. The participating teachers' attitudes, desires, and actions in this study toward improving their teaching behaviors serve as strong support for this conclusion.

Revisions to the process, based on suggestions for improvements offered by teachers, should result in an even more useful set of procedures. An improved student feedback process will ensue from measures taken to increase the feasibility of the process and to revise the method of compiling the results. Teacher involvement in tailoring this process at the local level will continue to be critical to their acceptance of the overall process.

One of the suggestions offered by teachers involved future administrations of the rating form by a neutral party. It is important to note that there may be negative consequences of this revision to the overall process on both the usefulness and feasibility for classroom teachers. If the "neutral" party happens to be an administrator from the same school, the teacher potentially puts herself at risk by having her formative feedback undergo possible review by an individual responsible for summative decision making. Similarly, if the "neutral" party is a colleague or peer, she again puts herself at risk by having someone other than herself read and review comments and ratings about her teaching. As a result of the potential review of data and results by the "neutral" party, teachers may be less willing to participate, fearing that they would be scrutinized or even judged by the third party.

Also, during the course of preparing the teachers' feedback, several open-ended comments were censored by the researcher, due to their harsh nature or the inclusion of profanity. Some teachers suggested that these comments should not be censored since the researcher could not have known the specific context being referred to in the comment. If comments of this nature were *not* censored, and teachers received all of the feedback, they might again be less willing to participate and open themselves up to feedback, especially if it included comments of this harsh nature. If teachers themselves or a neutral party were administering and analyzing the rating forms, the fact that comments would not be censored might again limit the usefulness of the process by discouraging teacher participation.

Additionally, it can be concluded that the *Student Evaluation of Teaching and Teaching Techniques* (*SE3T*) rating form is a highly useful and potentially valid instrument for the collection of student feedback of teaching performance. However, similar to the process developed for this study, the *SE3T* student rating form is in need of some minor revisions to specific items. As with the overall process, it is suggested that these revisions be made with a high degree of teacher involvement. Improvements to the rating form should be made and its usefulness, as well as its statistical characteristics, should continue to be assessed.

Finally, it is suggested that the content, strategy, and results of this study may be valuable resources to administrators and teachers. Administrators can utilize the outcomes as a means of motivating teachers within that school or district and encouraging them to use evaluation for self-improvement purposes. Through the application of this information, teachers can assess and improve their teaching performance according to the needs of their students. This, in and of itself, will enlighten teachers' attitudes toward the evaluation of their teaching performance.

A model of this process for collecting student feedback as a means of formative teacher evaluation was developed as a result of this study. *The Student Evaluation of Teaching and Teaching Techniques (SE3T) Model* is included as Appendix C.

A review of the literature indicates that few studies have been conducted which have investigated the perceptions held by teachers of the utility of receiving student feedback on

their teaching behaviors and performance. It is therefore recommended that additional studies be undertaken in order to determine the impressions of teachers (along with other indicators) regarding the usefulness of the process and subsequent feedback. Once there is a comprehensive understanding of teachers' perceptions and the impact of this form of feedback on the improvement of teaching behaviors, and ultimately on student performance, it may be possible to appropriately design teacher evaluation systems containing a student feedback component.

It is recommended that studies addressing the effect of student feedback on the improvement of teaching behaviors be undertaken. The goal of formative teacher evaluation is the improvement of instructional behaviors and practices. If a process of student feedback does not lead to improved teaching behaviors, then the implementation of that process has not served its general purpose. Teachers must not only be willing to *collect* this type of feedback, but must also be willing (and able) to *act* on the results. They must be willing to change their behaviors and must have access to the necessary resources. Demonstration of the effectiveness of student feedback on the improvement of teaching behaviors is requisite to the inclusion of such a process in a comprehensive system of teacher evaluation.

It is strongly suggested that the following recommendations with respect to the usefulness and feasibility of student evaluations of teaching be investigated. Specifically, it is recommended that

- ◊ the effect of student feedback on the improvement of teaching behaviors be examined;
- ◊ an assessment of the usefulness of the process in a variety of settings and with a variety of teachers, not just volunteers, be examined;
- ◊ the suggestions offered by teachers for improvements to the overall process should be incorporated in any ensuing student feedback studies;
- ◊ a possible alternative for the administration and subsequent guarantee of student anonymity, such as placing the *SE3T* rating instrument on-line in the school's computing network, be investigated
- ◊ students be given the opportunity to offer input into the development of items to be included on future versions of the *SE3T* rating instrument;
- ◊ steps be taken to educate and convince school administrators of the validity and usefulness of student feedback, specifically for the purpose of improving teacher performance; and
- ◊ research be conducted with teachers using this process for collecting student feedback in order to ascertain ways in which to structure professional development activities and inservice training.

Appendix A:

Frequencies and Percentages of Responses
for *Teacher Survey* Scaled Items

Table 1

Frequencies and Percentages of Responses for *Teacher Survey* Scaled Items

Item Description	Response Category	Frequency	Percentage
Usefulness of the process	Extremely useful	5	41.7
	Somewhat useful	7	58.3
	Not very useful	0	0.0
	Not at all useful	0	0.0
Appropriateness of *SE3T* items	Extremely appropriate	10	83.3
	Somewhat appropriate	2	16.7
	Not very appropriate	0	0.0
	Not at all appropriate	0	0.0
Feasibility for teachers	Yes	9	75.0
	No	2	16.7
	Unsure	1	8.3
Frequency of collection of student feedback	Once per school year	2	18.2
	Twice per school year	7	63.6
	Four times per school year	2	18.2
	More than four times	0	0.0

Appendix B:

Summary of Strengths and Concerns of the Process as
Identified in *Teacher Survey* and Focus Group

Table 2

Summary of Strengths and Concerns of the Process as Identified in *Teacher Survey* and Focus Group

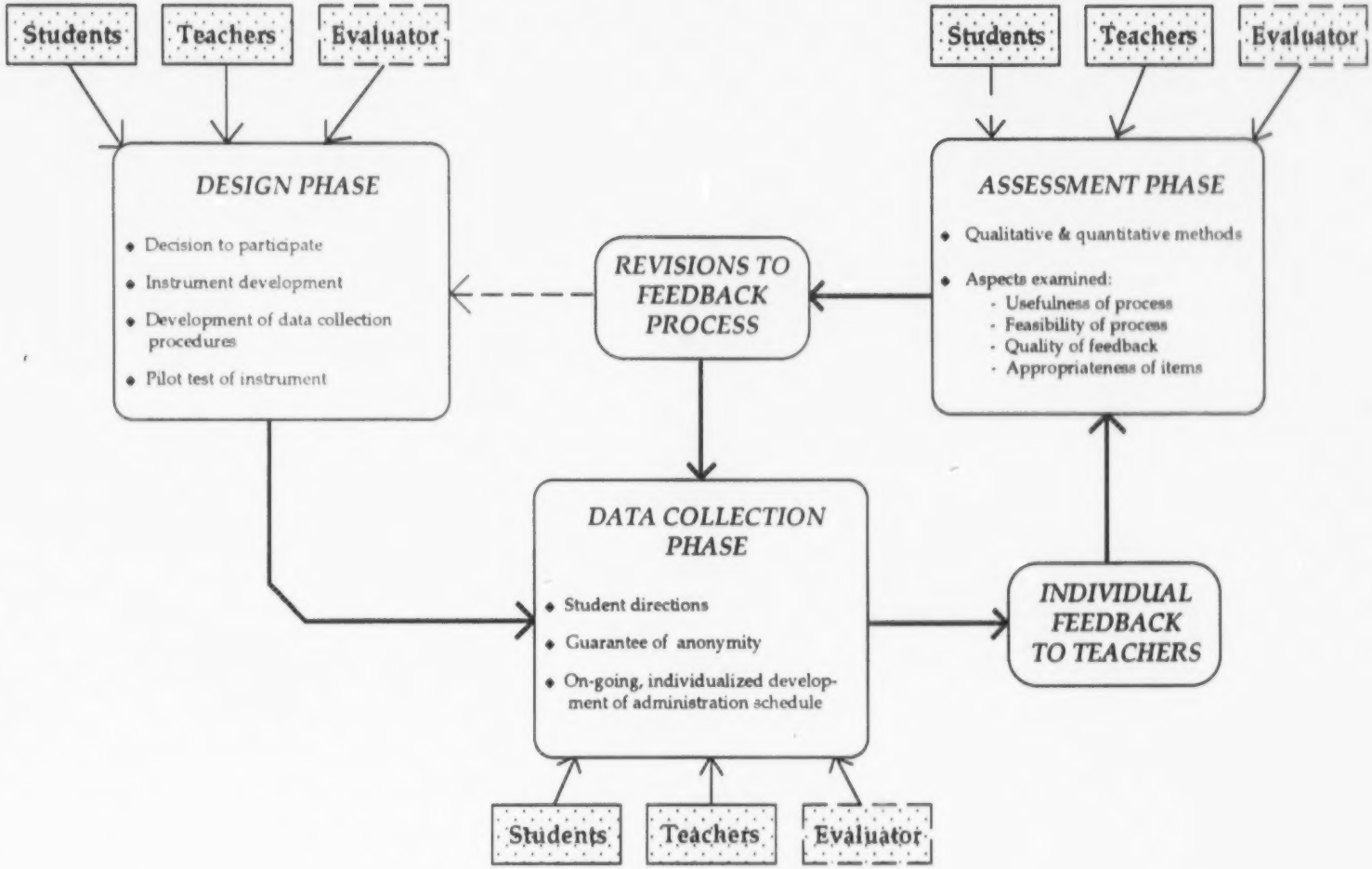
Characteristic	Strength	Concern
Usefulness of the process	• Allowed teachers to see themselves from another point of view • Was seen as important when working with a new group or level of students • Revealed areas of weakness early in professional career • Ensured higher level of honesty through anonymity	• May be decreased due to a lack of student honesty or failure of students to take the evaluation seriously • Requires ample time to process feedback prior to next administration • Is valuable only if teachers are willing to make changes
Appropriateness of *SE3T* items	• Covered a broad range of teacher behaviors • Was comprised of scaled and open-ended items that were equally useful • Consisted of sections that offered unique contributions to the feedback • Was supported for future use by teachers	• Included four items (13, 25, 30, and 33) identified by teachers as problematic
Feasibility of the process	• Involved a reasonable time frame for administering the form	• Might require a "neutral" proctor to protect student anonymity • Requires time and means for compiling results • Necessitates adapting collection procedures for certain groups of students
Frequency of *SE3T* administration	• Can be administered multiple times during a course	• Fails to provide enough time to improve teaching between frequent administrations

Table 2 – – continued.

Characteristic	Strength	Concern
Nature of teacher feedback	• Much value placed on responses to open-ended items	• Was difficult to discern if comments are personal or constructive • Would be enhanced by including all comments, due to specific contexts • Would be enhanced by reporting the distribution of ratings across all points
Changes in teacher behavior	• Identified and addressed several teacher behaviors including teacher empathy, verbal and nonverbal communication, pre-instructional planning, and varying methods	• Expressed an unwillingness to change behaviors
Teacher follow-up to the process	• Discussed results with students in open forum • Requested suggestions from students regarding how to improve specific behaviors • Used rating experience as a prompt for writing activities	• None identified

Appendix C:

Student Evaluation of Teaching and Teaching Techniques (SE3T) Model



BEST COPY AVAILABLE

REFERENCES

Aleomoni, L. M. & Spencer, R. E. (1973). The Illinois Course Evaluation Questionnaire: A description of its development and a report of some of its results. Educational and Psychological Measurement, 33, 669-684.

Bacharach, S. B., Conley, S. C., & Shedd, J. B. (1990). Evaluating teachers for career awards and merit pay. In J. Millman & L. Darling-Hammond (Eds.), The new handbook of teacher evaluation: Assessing elementary and

Bailey, G. D. (1978). Improving classroom instruction: Is there a better model? NASSP Bulletin, 62(414), 52-59.

Barber, L. W. (1990). Self-assessment. In J. Millman & L. Darling-Hammond (Eds.), The new handbook of teacher evaluation: Assessing elementary and secondary school teachers (pp. 216-228). Newbury Park, CA: Corwin.

Barsalou, J. M., Killinger, J. E., & Thompson, J. E. (1974). Student evaluation of staff in secondary schools. NASSP Bulletin, 58(379), 10-14.

Cook, S. S. (1989). Improving the quality of student ratings of instruction: A look at two strategies. Research in Higher Education, 30(1), 31-45.

Cangelosi, J. S. (1991). Evaluating classroom instruction. New York: Longman.

Darling-Hammond, L. (1986). A proposal for evaluation in the teaching profession. The Elementary School Journal, 86(4), 531-551.

Darling-Hammond, L., Wise, A. E., & Pease, S. R. (1983). Teacher evaluation in the organizational context: A review of the literature. Review of Educational Research, 53(3), 285-328.

Gitlin, A. & Smyth, J. (1990). Toward educative reform of teacher evaluation. Educational Theory, 40(1), 83-94.

Haefele, D. L. (1980). 'How to evaluate thee, teacher — Let me count the ways.' Phi Delta Kappan, 61(5), 349-352.

Haefele, D. L. (1992). Evaluating teachers: An alternative model. Journal of Personnel Evaluation in Education, 5(4), 335-345.

Haefele, D. L. (1993). Evaluating teachers: A call for change. Journal of Personnel Evaluation in Education, 7(1), 21-31.

Halbert, S. J. (1975). Student evaluation of teachers and administrators. Oregon School Study Council. (ERIC Document Reproduction Service No. 106 927)

Hanna, G. S., Hoyt, D. P., & Aubrecht, J. D. (1983). Identifying and adjusting for biases in student evaluations of instruction: Implications for validity. Educational and Psychological Measurement, 43(4), 1175-1185.

Huddle, G. (1985). Teacher evaluation — How important for effective schools? Eight messages from research. NASSP Bulletin, 69(479), 58-63.

Frase, L. E. & Streshly, W. (1994). Lack of accuracy, feedback, and commitment in teacher evaluation. Journal of Personnel Evaluation in Education, 8(1), 47-57.

Jacobson, P. O. (1986). The I-STEP teacher evaluation program for common schools. Illinois: Lincoln Land Community College. (ERIC Document Reproduction Service No. 268 658)

Johnston, J. M. & Hodge, R. L. (1981). Self-evaluation through performance statements — A basis for professional development. Journal of Teacher Education, 32(6), 30-33.

Joint Committee on Standards for Educational Evaluation. (1988). The personnel evaluation standards: How to assess systems for evaluating educators. Newbury Park, CA: Sage.

Levin, B. (1979). Teacher evaluation—A review of research. Educational Leadership, 37(3), 240-245.

Manning, R. C. (1988). The teacher evaluation handbook: Step-by-step techniques and forms for improving instruction. Englewood Cliffs, NJ: Prentice Hall.

Martin, D. V. (1988, February). Effective vs. less effective student teachers: How pupils rate their student teachers. Paper presented at the National Conference of the Association of Teacher Educators, San Diego, CA. (ERIC Document Reproduction Service No. 290 725)

McColskey, W. & Egelson, P. (1993). Designing teacher evaluation systems that support professional growth. Greensboro: University of North Carolina at Greensboro, SouthEastern Regional Vision for Education.

Nevo, D. (1994). How can teachers benefit from teacher evaluation. Journal of Personnel Evaluation in Education, 8(2), 109-117.

New York State Education Department. (1986). The ideal teacher. Albany: New York Education Department, Student Affairs Task Force. (ERIC Document Reproduction Service No. 275 645)

Pennsylvania Department of Education. (1981). Student Advisory Board, 1980-81 report. Harrisburg, PA: Pennsylvania Department of Education, Student Advisory Board. (ERIC Document Reproduction Service No. 210 815)

Savage, J. G. (1982). Teacher evaluation without classroom observation. NASSP Bulletin, 66(458), 41-45.

Savage, T. V. & McCord, M. K. (1986, April). The use of student evaluation in the assessment of teacher competence. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA. (ERIC Document Reproduction Service No. 278 105)

Scriven, M. (1980). The evaluation of college teaching. National Council of States on Inservice Education. (ERIC Document Reproduction Service No. 203 729)

Scriven, M. (1990). Teacher selection. In J. Millman & L. Darling-Hammond (Eds.), The new handbook of teacher evaluation: Assessing elementary and secondary school teachers (pp. 76-103). Newbury Park, CA: Corwin.

Scriven, M. (1991). Evaluation thesaurus (4th ed.). Newbury Park, CA: Sage.

Smith, J. P. & Brown, T. J. (1976, April). Relationships among secondary students' evaluations of their courses and teachers and their general attitude toward school. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA. (ERIC Document Reproduction Service No. 129 867)

Soar, R. S., Medley, D. M., & Coker, H. (1983). Teacher evaluation: A critique of currently used methods. Phi Delta Kappan, 65(4), 239-246.

Stiggins, R. J. (1986). Teacher evaluation: Accountability and growth systems—Different purposes. NASSP Bulletin, 70(490), 51-58.

Stiggins, R. J. & Duke, D. (1988). The case for commitment to teacher growth: Research on teacher evaluation. Albany: State University of New York.

Tacke, G. & Hofer, M. (1979). Behavioral changes in teachers as a function of student feedback: A case for the achievement motivation theory? Journal of School Psychology, 17(2), 172-180.

Traugh, C. E. & Duell, O. K. (1980). Secondary students' attitudes toward and experience with evaluating teachers. The High School Journal, 64(3), 97-107.

Veir, C. (1990). A tale of two appraisal systems: Texas and North Carolina. Unpublished manuscript, University of Texas at Austin. (ERIC Document Reproduction Service No. 327 981)